

What Is Claimed Is:

- 1 1. A method for classifying chromosome strings using a knowledge discovery engine,
2 comprising the steps of:
 - 3 a. processing the chromosome strings against a predefined set of
4 chromosome variables using an evolutionary computing algorithm, thereby
5 creating one or more processed chromosome strings;
 - 6 b. creating a lead cluster map for each processed chromosome string, wherein
7 each said lead cluster map comprises one or more clusters;
 - 8 c. computing a variance across all said clusters contained in each said lead
9 cluster map;
 - 10 d. determining a best lead cluster map, wherein said best lead cluster map is
11 said lead cluster map being the most homogeneous with said predefined set of
12 chromosome variables;
 - 13 e. determining whether said best lead cluster map is less than an acceptable
14 minimum; and
 - 15 f. re-processing said processed chromosome strings using said evolutionary
16 computing algorithm and repeating steps (b) through (f) if it is determined in said
17 step (e) that said best lead cluster map is not less than said acceptable minimum.
- 1 2. The method according to claim 1, further comprising:
 - 2 g. displaying said best lead cluster map if it is determined in said step (e) that
3 said best cluster map is less than said acceptable minimum.
- 1 3. The method according to claim 1, wherein a user inputs said acceptable minimum.
- 1 4. The method according to claim 1, wherein a user inputs said predefined set of
2 chromosome variables.

S&J Ref: 411520/00001

- 1 5. The method according to claim 1, wherein said evolutionary computing algorithm is a
2 genetic algorithm.
- 1 6. The method according to claim 1, wherein said predefined set of chromosome variables is
2 a set of bio-markers.
- 1 7. A knowledge discovery engine for classifying chromosome strings, comprising:
2 a means for evolutionary computing of the chromosome strings against a
3 predefined set of chromosome variables, thereby creating one or more processed
4 chromosome strings;
5 a means for pattern matching using an adaptive pattern recognition system for
6 organizing said processed chromosome strings;
7 a means for determining homogeneity of said processed chromosome strings with
 said predefined set of chromosome variables; and
9 a means for determining a best fit of said processed chromosome strings with said
10 predefined set of chromosome variables.
- 1 8. A computer program product for use with a computer system, comprising:
2 a computer usable medium having computer readable program code means
3 embodied in said medium for classifying chromosome strings using a knowledge
4 discovery engine, said computer readable program code means comprising:
5 processing means for enabling a processor to process the chromosome strings
6 against a predefined set of chromosome variables using an evolutionary computing
7 algorithm, thereby creating one or more processed chromosome strings;
8 creating means for enabling a processor to create a lead cluster map for each
9 processed chromosome string, wherein each said lead cluster map comprises one or more
10 clusters;

S&J Ref: 411520/00001

12

11 computing means for enabling a processor to compute a variance across all said
12 clusters contained in each said lead cluster map;

13 determining means for enabling a processor to determine a best lead cluster map,
14 wherein said best lead cluster map is said lead cluster map being the most homogeneous
15 with said predefined set of chromosome variables;

16 second determining means for enabling a processor to determine whether said best
17 lead cluster map is less than an acceptable minimum; and

18 re-processing means for enabling a processor to re-process said processed
19 chromosome strings using said evolutionary computing algorithm and repeating steps (b)
20 through (f) if it is determined in said step (e) that said best lead cluster map is not less
21 than said acceptable minimum.

S&J Ref: 411520/00001